Section 1: Introduction

This technical application relates to chapter 4. It is based on firm-level data for some unspecified country. It examines the impact of SPS and TBT measures implemented by the European Union (14 countries, Luxembourg is missing) on this country firms' exports. SPS and TBT data are extracted for the CEPII NTM dataset used in application 1.

Firm data are available for various years but NTM data are cross-section data. We therefore focus on the year 2011 and run cross-section estimations.

Both the firm-level data and NTM data are available on the course website.

The programme below finalizes the construction of the dataset by merging firm and NTM data.

Section 2: Data cleaning

We first restrict our sample of firms' exports to the year 2011 and clean the data by dropping incorrect product codes and domestic trade. We also keep only export flows to the European Union. Finally, we expand the dataset of firm exports and generate zero export flows. To do so, we create one European destination for each firm-product observation. In other words, we assume that for each product exported by a given firm, the flow can be potentially observed for each European destination. If it is not observed, then it is a zero flow with a value and a quantity set to zero. For non-zero observations, we merge our observations with the initial firm data in order to get the value and the quantity exported by the firm.

Section 3: Data manipulations

Before uploading data we have to increase the maximum number of variables that can be included in any of Stata's estimation commands. The default size is 400. However, with non-linear estimation fixed effects will have to be included using dummy variables. With firm level data the number of dummy variables is likely to be larger than 400. We then set this number to 4000.

```
set matsize 4000
```

We can now upload our data and proceed with data cleaning.

Before opening the dataset you may want to change the default directory that is the directory STATA will use unless another directory is specified. We created a folder called STATA_NTM (c:/ STATA_NTM) and copied all files that are needed to run the application (data files and do-files) into that folder. To change the default directory just run

```
cd c:/ STATA_NTM/Apllication3
```

```
use FIRMS.dta, clear
```

```
* If you had not changed the default directory the right command would be
* use c:/ STATA_NTM/ FIRMS.dta, clear


rename y year
```

For many countries, NTMs in the CEPII dataset (version 2014) are for 2011. We decide to work with the 2011 exports of firms exclusively

```
keep if year == 2011
```

A first glance at the data allows us to identify some mistakes in HS products coding. The hs variable is a non-numerical or string variable. We keep only those codes which are 6 digit long. We then drop any incorrectly specified HS product code

```
gen l = length(hs)

tab l

drop if l != 6

drop l

rename hs hs6

rename d ccode_d
```

*We observe some domestic trade (the d variable is equal to "XXX") that may simply be a reporting error. It is safer to drop such observations

```
drop if ccode_d == "XXX"
```

We rename some variables in order to make the name more informative.

```
rename f firm

order firm year hs6 ccode_d
```

It is important to get an idea of the final size of the sample. We thus run the count command

```
count
```

* As we want to focus on export flows to EU15 we have to construct a dummy variable identifying EU15 (minus Luxembourg) countries

```
gen eu15_d = 1 if ccode_d == "AUT" | ccode_d == "BEL" | ccode_d == "DEU" | ccode_d =="DNK" | ccode_d == "ESP" | ccode_d == "FIN" | ccode_d =="FRA" | ccode_d == "GBR" | ccode_d == "GRC" | ccode_d == "IRL" | ccode_d == "ITA" | ccode_d == "NLD" | ccode_d == "PRT" | ccode_d == "SWE"
```

```
replace eu15_d = 0 if eu15_d == .
```

We then remove any other destination.

```
keep if eu15_d == 1
```

```
sort firm hs6 ccode_d
```

Data are sorted as we will have to merge datasets later on. Note than sorting data is not necessary with STATA 12 and later STATA versions.

```
save temp_firmdata, replace
```

Section 4: Data manipulations (continued)

We now expand the dataset in order to create for each firm-product combination, all destinations (i.e. the potential number of destinations: 14 EU countries - LUX is missing). This is done in several steps.

We first keep one observation per firm-hs6 code

```
keep firm hs6
```

```
bys firm hs6 : gen obs = _n
```

```
keep if obs == 1
```

```
drop obs
```

```
codebook hs6, compact
```

```
codebook firm, compact
```

We then create an indicator named fp identifying uniquely each firm product combination using the egen command associated with the group function.

```
egen fp = group(firm hs6)
```

for each fp, that is each firm-product combination, we create 14 observations: one for each potential EU destination using the expand command.

```
expand 14:
```

We can then associate each observation with a destination taken from our EU14 group.

```
bys fp : gen obs_expand = _n
```

```
gen ccode_d = "AUT" if obs_expand == 1
```

```
replace ccode_d = "DEU" if obs_expand == 2

replace ccode_d = "DNK" if obs_expand == 3

replace ccode_d = "FIN" if obs_expand == 4

replace ccode_d = "FRA" if obs_expand == 5

replace ccode_d = "GBR" if obs_expand == 6

replace ccode_d = "GRC" if obs_expand == 7

replace ccode_d = "IRL" if obs_expand == 8

replace ccode_d = "ITA" if obs_expand == 9

replace ccode_d = "NLD" if obs_expand == 10

replace ccode_d = "PRT" if obs_expand == 11

replace ccode_d = "SWE" if obs_expand == 12

replace ccode_d = "ESP" if obs_expand == 13

replace ccode_d = "BEL" if obs_expand == 14
```

We can now remove redundant variables and move to our merge operation.

```
drop obs_expand fp

sort firm hs6 ccode_d

save temp_firmdata_expand, replace
```

Section 5: Merging datasets

We merge our dataset with the temp_firmdata dataset saved previously in order to get information on both exported quantity and value for strictly positive flows. All missing values for both export quantity and export value variables will be in fact zeros.

```
merge firm hs6 ccode_d using temp_firmdata

tab _merge

drop _merge

replace year = 2011 if year == .

replace v = 0 if v == .
```

```
replace q = 0 if q == .

save database, replace
```

We can always save disk space by erasing some superfluous files.

```
erase temp_firmdata.dta

erase temp_firmdata_expand.dta
```

Section 6: NTM data

We can now deal with NTM data. We first restrict our data to SPS and TBT measures implemented by European countries. We then merge firm data with NTM data and define two NTM variables: a simple dummy "Pres_SPSTBT" set to one if the European Union applies at least one SPS or TBT on a given HS 6-digit product (0 otherwise). The second NTM variable ("num_SPSTBT") represents the total number of SPS and TBT measures applied by the European Union on a given product.

Let's upload the NTm data first.

Section 7: NTM data (continued)

```
use Base_Mast_Oct2014.dta, clear
```

As the focus on SPS and TBTs we keep only relevant variables.

```
keep isor product numA numB PresA PresB

rename isor ccode_d
```

As before we drop any destination that is not a EU15 country.

```
gen eu15_d = 1 if ccode_d == "AUT" | ccode_d == "BEL" | ccode_d == "DEU" | ccode_d == "DNK" | ccode_d == "ESP" | ccode_d == "FIN" | ccode_d == "FRA" | ccode_d == "GBR" | ccode_d == "GRC" | ccode_d == "IRL" | ccode_d == "ITA" | ccode_d == "NLD" | ccode_d == "PRT" | ccode_d =="SWE"

replace eu15_d = 0 if eu15_d == .

keep if eu15_d == 1

codebook ccode_d, compact

count
```

The product code variable is numeric. As in our firm dataset it is in a string format we make the necessary transformation to make then compatible with each other.

```
gen hs6 = string(product)
```

```
gen l = length(hs6)

tab l

replace hs6 = "0"+ hs6 if l == 5

drop product l

sort ccode_d hs6

save temp_NTM, replace
```

Section 8: Final dataset

We are ready now to merge the firm-level data with the NTM data

```
use database, clear

sort ccode_d hs6

merge ccode_d hs6 using temp_NTM

tab _merge
```

Our reference dataset is the firms' dataset as we are interested in the impact of NTMs on firms' exports. As a consequence we can get ride on any observation on NTMs that does not apply to any export flow in the firms' dataset.

```
drop if _merge == 2

drop _merge
```

Wherever no NTM applies NTM variables will have a missing value. We replace this missing occurrences by zeros.

```
replace numA = 0 if numA == .

replace numB = 0 if numB == .

replace PresA = 0 if PresA == .

replace PresB = 0 if PresB == .
```

We can now define our two variables for SPS and TBTs presence and incidence. Remember we treat SPS measures and TBTs indistinctly.

The first variable is a simple dummy variable that takes value 1 wherever at least one NTM is applied

```
gen Pres_SPSTBT = 1 if PresA == 1 | PresB == 1

replace Pres_SPSTBT = 0 if Pres_SPSTBT == .
```

The second NTM variable is the total number of SPS and TBTs observed at the product-importer level.

```
gen num_SPSTBT = numA + numB

drop PresA PresB numA numB

save database, replace

erase temp_NTM.dta
```

Section 9: Econometric Estimations

We are now ready to run some estimations.

We first consider the simple dummy NTM variable. To keep zero flows in our regressions, we use the Poisson estimator. The estimator treats dependent variables as if they were count variables. Zero observations can thus be included. The estimator has other important properties and in particular it is efficient in treating heteroscedasticity which particularly characterized trade flows.

Our estimations also include various sets of fixed effects: sectors fixed effects defined at HS 2-digit, firms' fixed effects, and importing country fixed effects to control for all unobservable characteristics at the sector, firm and destination levels that may influence the bilateral exports of firms.

Note that our sample contains just on exporting country and therefore the importer fixed effect is equivalent to a country-pair fixed effect. This means that all bilateral geographical variables do not have to be included as their effect is already accounted for.

Section 10: Econometric Estimations (continued)

In this first set of estimations, we just include a simple dummy for NTMs controlling for the presence versus the absence of SPS or TBT measures at the product-destination level.

We consider three different dependent variables: the value of the export flow (regressions (1) and (2)), the exported volume (regressions (3) and (4)), and the unit value of the flow (regression (5)). Regressions (1) and (3) include all flows, while regressions (2), (4) and (5) are restricted to strictly positive export flows.

Results as they appear in the main STATA windows can be saved in some TXT file. This is done by using the log command. You must open your text file and then close with the log close command wherever you are not interested by the outcome anymore. There is also the possibility to export your

estimated coefficients in a word document in table format. This is possible by using the esttab command that reports all results saved with the eststo command.

```
log using results.txt, text replace

use database, clear
```

Section 11: Econometric Estimations (continued)

As mentioned previously as we are using a non-linear estimator fixed effects will have to be introduced explicitly with the help of dummy variables. We then generate an all set of sector specific dummies at the HS 2 digit level. We thus substract the first two digit of our 6 digit HS code.

```
gen hs2 = substr(hs6, 1 , 2)
```

As mentioned previously we start with the value of exports as dependent variable.

In Regression #1 all flows are considered. Note that except for our NTM variable all control variables are fixed effects related.

```
xi: poisson v Pres_SPSTBT i.hs2 i.ccode_d i.firm, difficult robust
```

In Regression #2 only strictly positive exported values are considered.

In this case we also run an estimation with a linear estimator and taking as our dependent variable the natural log of exports values.

```
xi: poisson v Pres_SPSTBT i.hs2 i.ccode_d i.firm if v > 0, difficult robust

eststo

gen lnv=ln(v)

xi: reg lnv Pres_SPSTBT i.hs2 i.ccode_d i.firm if v > 0, robust

eststo
```

We then consider the quantity of exports as our dependent variable.

Regression #3 includes all flows

```
xi: poisson q Pres_SPSTBT i.hs2 i.ccode_d i.firm, difficult robust

eststo
```

Regression #4 includes only strictly positive exported volumes

```
xi: poisson q Pres_SPSTBT i.hs2 i.ccode_d i.firm if q > 0, difficult robust

eststo
```

Here again we estimate a linear in logs specification on strictly positive values.

```
gen lnq=ln(q)

xi: reg lnq Pres_SPSTBT i.hs2 i.ccode_d i.firm if q > 0, robust

eststo
```

Dependent variable: unit value of exports

Regression #5: only strictly positive unit values

```
gen uv = v / q

xi: poisson uv Pres_SPSTBT i.hs2 i.ccode_d i.firm if uv > 0, difficult robust

eststo
```

Again we estimate a linear in logs specification.

```
gen lnuv=ln(uv)

xi: reg lnuv Pres_SPSTBT i.hs2 i.ccode_d i.firm if uv > 0, robust

eststo
```

All our results are reported in a table saved in the NTM_REG_firm_2 file.

We will comment them at the very end of the econometric exercise.

```
esttab using NTM_REG_Firm_2.rtf, keep(Pres_SPSTBT) title(Presence of SPS/TBT) star(c 0.10 b 0.05 a 0.01) replace
```

Section 12: Econometric Estimations (continued)

We now re-run the estimations controlling for the number of NTMs. In other words, we replicate the first set of estimations, but instead of using a simple dummy for NTMs, we now include the total number (in logs) of SPS and TBT measures notified by European importers on each HS 6-digit product. We therefore deal only with importer-product observations for which this number is strictly positive.

Log number of SPS/TBT measures at the importer-product level

```
gen lnum_SPSTBT = ln(num_SPSTBT)

eststo clear
```

Dependent variable: value of exports

Regression #1: all flows

```
xi: poisson v lnum_SPSTBT i.hs2 i.ccode_d i.firm if num_SPSTBT > 0, difficult robust

eststo
```

Regression #2: only strictly positive exported values

```
xi: poisson v lnum_SPSTBT i.hs2 i.ccode_d i.firm if num_SPSTBT > 0 & v > 0, difficult robust

eststo

xi: reg lnv lnum_SPSTBT i.hs2 i.ccode_d i.firm if num_SPSTBT > 0 & v > 0,  robust

eststo
```

Dependent variable: quantity of exports

Regression #3: all flows

```
xi: poisson q lnum_SPSTBT i.hs2 i.ccode_d i.firm if num_SPSTBT > 0, difficult robust

eststo
```

Regression #4: only strictly positive exported volumes

```
xi: poisson q lnum_SPSTBT i.hs2 i.ccode_d i.firm if num_SPSTBT > 0 & q >0, difficult robust

eststo

xi: reg lnq lnum_SPSTBT i.hs2 i.ccode_d i.firm if num_SPSTBT > 0 & q > 0,  robust

eststo
```

Dependent variable: unit value of exports

Regression #5: only strictly positive unit values

```
xi: poisson uv lnum_SPSTBT i.hs2 i.ccode_d i.firm if num_SPSTBT > 0 & uv > 0, difficult robust

eststo

xi: reg lnuv lnum_SPSTBT i.hs2 i.ccode_d i.firm if num_SPSTBT > 0 & uv > 0,  robust
```

```
eststo

esttab using NTM_REG_Firm_2.rtf, keep(lnum_SPSTBT) title(Number of SPS/TBT) star(c 0.10 b 0.05 a 0.01) append

log close
```

We can now open our rtf file in any standard documents editor software.

Presence of SPS/TBT

	(1) v	(2) v	(3) lnv	(4) q	(5) q	(6) lnq	(7) uv	(8) lnuv
main								
Pres_SPSTBT	-1.271b	-1.652a	-0.0675	-2.262a	-1.410a	-0.231	0.161	0.164
	(-2.45)	(-5.02)	(-0.21)	(-4.30)	(-8.54)	(-0.86)	(0.50)	(0.74)
N	18116	1660	1660	18116	1659	1659	1659	1659

t statistics in parentheses
c $p < 0.10$, b $p < 0.05$, a $p < 0.01$

Number of SPS/TBT

	(1) v	(2) v	(3) lnv	(4) q	(5) q	(6) lnq	(7) uv	(8) lnuv
main								
lnum_SPSTBT	1.774c	1.513c	0.0895	-1.514a	-1.430a	-0.0279	0.130	0.125
	(1.70)	(1.71)	(0.38)	(-2.62)	(-3.38)	(-0.10)	(0.43)	(0.65)
N	14900	1380	1380	14900	1379	1379	1379	1379

t statistics in parentheses
c $p < 0.10$, b $p < 0.05$, a $p < 0.01$

Section 13: Results

We can first observe that standard linear estimations do not show any significant results. When focusing on Poisson estimations, significantly different from zero coefficients are obtained with value and quantity of exports as dependent variables. Results are consistent across NTM indicator only for regressions with exports quantity as dependent variable. The impact of NTMs on exports value is negative when the presence of NTMs is measured by a dummy and positive when measured by the number of applied NTMs. It is not obvious to find any convincing explanation for such contrasting results.